

# Diagnostic performance of the PTSD checklist and the Vietnam Era Twin Registry PTSD scale

K. Magruder<sup>1,2\*</sup>, D. Yeager<sup>1,2</sup>, J. Goldberg<sup>3,4</sup>, C. Forsberg<sup>3</sup>, B. Litz<sup>5,6</sup>, V. Vaccarino<sup>7</sup>, M. Friedman<sup>8,9</sup>,  
T. Gleason<sup>10</sup>, G. Huang<sup>10</sup> and N. Smith<sup>3,4</sup>

<sup>1</sup> Mental Health Service, Ralph H. Johnson VA Medical Center, Charleston, South Carolina, USA

<sup>2</sup> Department of Psychiatry, Medical University of South Carolina, Charleston, South Carolina, USA

<sup>3</sup> Seattle Epidemiologic Research and Information Center, VA Puget Sound Health Care System, Seattle, Washington, USA

<sup>4</sup> Department of Epidemiology, University of Washington, Seattle, Washington, USA

<sup>5</sup> Massachusetts Epidemiology Research and Information Center, VA Boston Healthcare System, Boston, Massachusetts, USA

<sup>6</sup> Boston University School of Medicine, Boston, Massachusetts, USA

<sup>7</sup> Departments of Epidemiology and Medicine, Emory University, Atlanta, Georgia, USA

<sup>8</sup> Department of Veterans Affairs, National Center for Posttraumatic Stress Disorder, White River Junction, Vermont, USA

<sup>9</sup> Departments of Psychiatry and Pharmacology & Toxicology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

<sup>10</sup> Cooperative Studies Program, Clinical Science Research and Development, VA Office of Research and Development, Washington, DC, USA

**Aims.** Self-report questionnaires are frequently used in clinical and epidemiologic studies to assess post-traumatic stress disorder (PTSD). A number of studies have evaluated these scales relative to clinician administered structured interviews; however, there has been no formal evaluation of their performance relative to non-clinician administered epidemiologic assessments such as the Composite International Diagnostic Interview (CIDI). We examined the diagnostic performance of two self-report PTSD scales, the PTSD checklist (PCL) and the Vietnam Era Twin Registry (VET-R) PTSD scale, compared to the CIDI.

**Methods.** Data were derived from a large epidemiologic follow-up study of PTSD in 5141 Vietnam Era Veterans. Measures included the PCL, VET-R PTSD scale and CIDI. For both the PCL and VET-R PTSD scale, ROC curves, areas under the curve (AUC), sensitivity, specificity, % correctly classified, likelihood ratios, predictive values and quality estimates were generated based on the CIDI PTSD diagnosis.

**Results.** For the PCL and VET-R PTSD scale the AUCs were 89.0 and 87.7%, respectively. Optimal PCL cutpoints varied from the 31–33 range (when considering sensitivity and specificity) to the 36–56 range (when considering quality estimates). Similar variations were found for the VET-R PTSD, ranging from 31 (when considering sensitivity and specificity) to the 37–42 range (when considering quality estimates).

**Conclusions.** The PCL and VET-R PTSD scale performed similarly using a CIDI PTSD diagnosis as the criterion. There was a range of acceptable cutpoints, depending on the metric used, but most metrics suggested a lower PCL cutpoint than in previous studies in Veteran populations.

Received 23 December 2013; Revised 14 April 2014; Accepted 16 April 2014; First published online 6 June 2014

**Key words:** Assessment, CIDI, PCL, PTSD.

## Introduction

Post-traumatic stress disorder (PTSD) is a common psychiatric disorder with lifetime prevalence in the general US population of 3.6% for males and 9.7% for females (Kessler *et al.* 2005). There are several different ways to formally diagnose PTSD. Diagnostic assessments range from structured interviews administered by specially trained clinicians, such as the

PTSD module of the Structured Clinical Interview for Diagnostic and Statistical Manual (DSM-IV) (Spitzer *et al.* 1990) (SCID) or the Clinician Administered PTSD Scale (CAPS) (Blake *et al.* 1995), to structured non-clinician administered instruments, such as the Diagnostic Interview Schedule (DIS) (Robins *et al.* 1981) or the Composite International Diagnostic Interview (CIDI) (Robins *et al.* 1988). In many large epidemiological studies of PTSD, it is not practical or feasible to use trained interviewers. Instead, self-report screening questionnaires are commonly used to assess PTSD symptom severity burden and to generate a presumptive diagnosis. [See, for example (Schuster *et al.* 2001; Schlenger *et al.* 2002;

\* Address for correspondence: Dr K. M. Magruder, Mental Health Service, Ralph H. Johnson VA Medical Center, 109 Bee Street, Charleston, SC 29401, USA.

(Email: [magrudkm@musc.edu](mailto:magrudkm@musc.edu))

Hoge *et al.* 2004). By far the most commonly used self-report PTSD questionnaire is the PTSD checklist (PCL) (Elhai *et al.* 2005; McDonald & Calhoun, 2010). Developed in 1993 (Weathers *et al.* 1993), the PCL consists of 17 items corresponding to the symptoms specified in the DSM, 4th edition (American Psychiatric Association, 1994).

Two studies (Lang *et al.* 2003; Lang & Stein, 2005) have examined the PCL relative to the CIDI, both in Veterans Health Administration primary care settings. The performance characteristics of the PCL were not markedly different from those of other studies that compared the PCL to clinically based diagnostic assessments. However, these previous studies were based on clinical samples, and there are no studies comparing the performance of the PCL to the CIDI in a community sample.

*The Course and Consequences of PTSD in Vietnam-Era Veteran Twins* (VA Cooperative Study #569 or CSP #569) is an investigation of PTSD in Veterans in the Vietnam-Era Twin Registry (VET-R). The PTSD assessment protocol for this study included the CIDI PTSD module and the PCL. To capture longitudinal change in PTSD symptom burden the protocol also used a 15-item PTSD scale (Goldberg *et al.* 1990; True *et al.* 1993; Boscarino, 2006, 2008) that had been previously administered in 1985–1987 when the Registry was created (Roy-Byrne *et al.* 2004). This VET-R PTSD scale pre-dated the development of the PCL and has been used extensively in VET-R research projects, thus its relationship with new and more widely used assessments will help to interpret historical data that are important in characterizing this cohort.

The aim of this paper was to evaluate the diagnostic performance of the PCL and the VET-R PTSD scale compared with a CIDI PTSD diagnosis. Understanding the performance of these instruments in comparison with the CIDI is of interest to future studies using this combination of measures and will also enable comparisons with large epidemiologic studies (e.g., World Mental Health Survey) when it is not practical to administer the CIDI.

## Methods

### Study design and setting

#### Design

CSP #569 is an observational study of PTSD among Veterans who are members of the VET-R (Henderson *et al.* 1990). The study collected data by mail and telephone between January 2010 and September 2012. A mail questionnaire obtained general health information and PTSD symptoms with the PCL and VET-R

PTSD scale. A telephone interview used the CIDI to diagnose PTSD.

#### Setting

The VET-R is the source of Vietnam-Era Veterans for this study. It was constructed from computerised military discharge records and is a national sample of male twins from all branches who served on active duty during the Vietnam Era (1964–1975) (Eisen *et al.* 1987; Henderson *et al.* 1990; Goldberg *et al.* 2002). Members of the Registry were born between 1939 and 1957 representing a cross-section of Veterans of the era who were assembled on the basis of being members of a twin pair. The VET-R is a non-clinical sample and was not compiled based on post-discharge medical or psychiatric disorders or health seeking behaviours. A complete description of the VET-R has been published (Eisen *et al.* 1987; Henderson *et al.* 1990; Goldberg *et al.* 2002).

## Subjects and methods

### Subjects

All members of the VET-R who were known to be alive and had not withdrawn were recruited to participate in this study. Informed consent was obtained from all participating VET-R members, and the VA Central Institutional Review Board approved the study's protocol.

### Field work

We mailed all twins an initial letter with a 20-page health survey and invited them to participate. Twins were instructed to complete and return the survey by mail. For twins who did not send back a survey, we attempted to call them directly. Because of the size and scope of the study, all mail and telephone fieldwork was done under contract by Abt-SRBI, Inc.

## Measures

### PTSD diagnosis

PTSD was assessed by a telephone administration of the CIDI according to the 4th edition of the DSM-IV (American Psychiatric Association, 1994). The CIDI is a structured instrument administered by trained lay interviewers (Robins *et al.* 1988). Interview training was done by CIDI certified trainers at the start of the study and continuously monitored during the course of the fieldwork. We assessed each respondent for current PTSD (an episode within the 12 months prior to the interview).

### Dimensional assessment of PTSD

The mailed survey included the PCL (Weathers *et al.* 1993). As VET-R members were discharged from the military decades ago, we used the civilian, rather than the military, version of the PCL (Weathers & Ford, 1996). For each item the respondent was asked to rate how bothered he has been in the past month by that symptom on a 5-point scale from 'not at all' (a value of 1) to 'extremely' (a value of 5). Responses to the 17 items were summed to generate the PCL score which can range from 17 to 85. Developed in 1993 by researchers at the National Center of PTSD, initial psychometric data showed that the PCL correlates highly ( $r=0.93$ ) with the CAPS (Blake *et al.* 1990), has good diagnostic efficiency ( $>0.70$ ) and robust psychometric properties with a variety of traumatically exposed populations (Blanchard *et al.* 1996; Andrykowski *et al.* 1998; Manne *et al.* 1998). A score of 50 was originally recommended for a presumptive diagnosis of PTSD; however, in recent years analyses support the use of lower cutoff scores (Yeager *et al.* 2007; Yeager & Magruder, 2013).

The mailed survey also included the VET-R PTSD a 15 item-scale that assesses the frequency of symptoms in the previous 6 months. The items were derived based on the DSM-III-R (American Psychiatric Association, 1987) criteria and asked in the context of 'military service experiences' (see online Supplemental Material for the question items). Response options range from 1 'never' to 5 'very often' along a 5-point ordinal scale. The total VET-R PTSD score ranges from 15 to 75. The total scale demonstrated excellent internal consistency (Cronbach's  $\alpha=0.9$ ) and the test-retest reliability was 0.6 (Roy-Byrne *et al.* 2004). The rationale for readministering the VET-R PTSD scale was that it permits a direct comparison of the course of PTSD symptoms across a 25-year interval.

The two PTSD self-report measures were contained in separate (non-contiguous) sections of the mailed survey material. We considered randomly assigning the order of the scales within the survey, but because PCL uses the more contemporary DSM-IV criteria and was more relevant to the primary study question, we made that scale appear first in all questionnaires.

### Demographic and military service factors

Data on key demographic and military service characteristics were available from archival information in the VET-R. Variables used descriptively in the present study include age, race, branch of service, military rank and military service in the Southeast Asia theatre of operations.

### Statistical analysis

Descriptive statistics examined the mean or percent distribution for demographic and military service characteristics. For both scales, we calculated Cronbach's  $\alpha$  as an estimate of internal reliability; we also used the Pearson  $r$  to examine the correlation between the two scales. Using the CIDI PTSD diagnosis as the criterion, we calculated sensitivity and specificity at each point, as well as positive and negative predictive values and the percent correctly classified. Optimal cutoff points were identified that maximised the percent correctly classified.

We plotted the receiver operator characteristic (ROC) curve at every value of the PCL and VET-R PTSD scale. The ROC curve is a graphical plot of sensitivity *v.* one-specificity across all possible values on the PCL and the VET-R PTSD scale. ROC plots provide a visual way to examine the tradeoffs between sensitivity and specificity at any scale value. We also estimated the area under the curve (AUC) values for both PCL and the VET-R PTSD scale, which provides an index of the predictive accuracy of a measure. Typically an AUC value  $\geq 80\%$  is considered to be indicative of a scale with good diagnostic accuracy (Kraemer, 1987). We adjusted the AUC confidence intervals for clustering due to the sample of twin pairs.

We also calculated positive and negative likelihood ratios (LR+ and LR-), which are based on the ratio of sensitivity and specificity and are unaffected by prevalence. Larger values for these ratios indicate a higher probability of having PTSD and conversely, smaller values correspond to a lower probability of having PTSD. LR+ values that exceed 5 and LR- values that are below 0.2 are considered informative for the likelihood of the presence or absence of the disorder. Last, we calculated quality estimates for  $k(0.0)$ ,  $k(1.0)$ , and  $k(0.5)$ . Kappa ( $k$ ) is a measure of diagnostic efficiency which is often evaluated at weights of 0.0 (optimising sensitivity), 1.0 (optimising specificity) and 0.5 (optimising efficiency) (Kraemer, 1987). The target range for  $k(0.5)$  is between 0.4 and 0.6; however, 0.2–0.4 is considered acceptable, and 0.6–0.8 'a cause for celebration' (Kraemer *et al.* 2012).

All analyses were done in Stata 11.2 (StataCorp, 2009).

## Results

### Sample characteristics

In total 6692 twins responded to the mail survey with valid PCL and VET-R PTSD scale data. Of those, 5141 (76.8%) twins also completed the CIDI PTSD diagnostic module and comprised our analytic sample. The

mean age of the sample was 61.1 years and the majority were white (90.6%); most men had served in the Army (51.1%), were enlisted (96.2%), and 36.0% served in the Southeast Asia theatre of operations (Table 1).

### Psychometric characteristics

Cronbach's  $\alpha$  was 0.96 for both the PCL and the VET-R PTSD scale. The scales were highly correlated ( $r=0.9$ ). The mean PCL score for those with and without a current PTSD diagnosis (by CIDI) was significantly different (49.6 *v.* 25.2;  $p<0.001$ ). Likewise, the mean value of the VET-R PTSD scale was 44.3 for those with current PTSD and 23.9 for those without PTSD ( $p<0.001$ ).

**Table 1.** Demographic and military service factors

Demographic and military service factors	<i>n</i>	(%)
Age		
50–59	1383	26.9
60–61	1092	21.2
62–63	1656	32.2
64 and older	1010	19.7
Race		
White	4660	90.6
Non-white	431	8.4
Ethnicity*		
Hispanic	149	2.9
Non-hispanic	4780	93.0
Current marital status*		
Married/partnered/widowed	4119	80.1
Divorced/never married	1001	19.5
Current education		
Less than HighSchool graduate	227	4.4
High School graduate	1145	22.3
Some college/Vocational	2167	42.2
College graduate or more	1602	31.2
Current employment status*		
Full time	2309	44.9
Part time	625	12.2
Retired	1310	25.5
Disabled	446	8.7
Unemployed	426	8.3
Military rank*		
Enlisted	4946	96.2
Officer	186	3.6
Branch of service		
Army	2628	51.1
Navy	1233	24.0
Air force	939	18.3
Marines	341	6.6
Southeast Asia service		
Yes	1849	36.0
No	3292	64.0

\*Variable does not sum to 5141 due to missing information.

The AUC for the PCL was 89.0 and 87.7% for the VET-R PTSD scale; the ROC shapes were similar (Figs 1 and 2).

The ROC analysis comparing the PCL with the CIDI PTSD diagnosis found cutoff points from 31 to 33 to be optimal; in this range both sensitivity and specificity were  $>80$  and  $>80\%$  were correctly classified (Supplemental Table A). Optimal LR+ yielded cutoff points  $\geq 34$ , with optimal LR– cutoff points  $\leq 31$ . The predictive value positive ranged from 27.4 to 30.5% and predictive value negative from 98.0 to 98.4% for scores between 31 and 34. The optimal quality estimates for cutoff scores in this range were  $k(0.5)=33.8$ – $37.2\%$ , which is lower than the maximum  $k(0.5)$  of 45.3% corresponding to a PCL cutoff score of 50; PCL scores from 36 to 56 are considered in the target range, with scores from 24 to 35 still acceptable.

Based on a sensitivity of 82.2% and a specificity of 81.4%, 31 was the optimal cutpoint value for the VET-R PTSD scale (Supplemental Table B). Optimal LR+ yielded cutoff points  $\geq 34$ , with optimal LR– cutoff points  $\leq 28$ . The predictive value positive ranged from 24.5 to 29.8% and predictive value negative from 97.8 to 98.2% for scores between 29 and 33. The optimal quality estimates for cutoff scores in this range were  $k(0.5)=29.7$ – $36.0\%$ , lower than the maximum  $k(0.5)$  of 41.9% which corresponds with a VET-R cutoff score of 39; VET-R PTSD scale scores from 37 to 42 are considered in the target range, with scores from 24 to 36 still acceptable.

### Southeast Asia service

We examined the PCL and VET-R PTSD scales by Southeast Asia service (yes/no) to see if either one performed differentially for those with war-zone experience. When stratified by theatre service, the PCL AUCs were 90.6% (88.4–92.9%) for those who served in theatre and 86.9% (84.1–89.8%) for those not in theatre and were significantly different ( $p=0.04$ ). The difference between the service strata AUCs for the VET-R PTSD scale was also statistically significant (AUC in theatre = 90.0% (87.7–92.2%), AUC not in theatre = 85.0% (82.1–87.8%)) suggesting that the both scales performed slightly better among those who served in Southeast Asia. All AUC confidence intervals were adjusted for clustering within twin pairs.

### Discussion

Both the PCL and the VET-R PTSD scale demonstrated similarly excellent psychometric properties in our national study of older Vietnam-Era Veterans relative to a commonly used scale. While ours is not the first

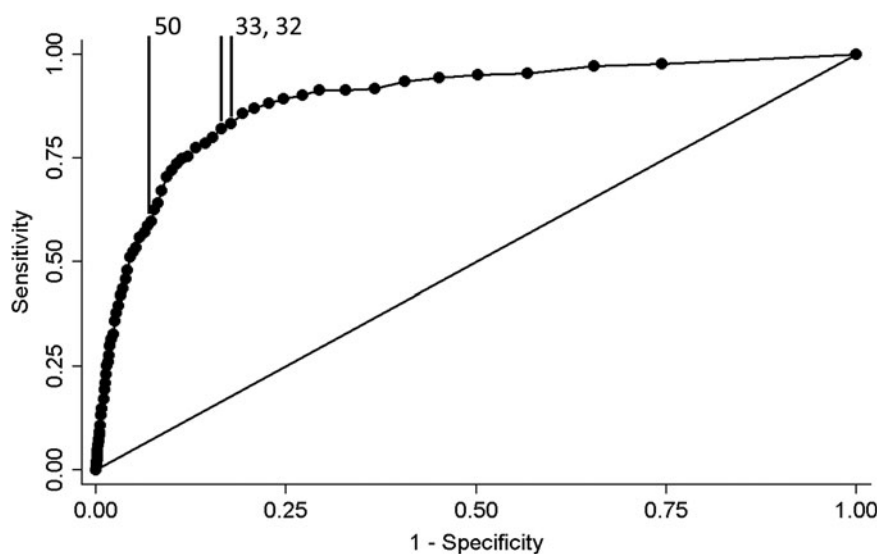


Fig. 1. ROC curve for the PCL.

study to compare a PTSD self-report questionnaire, such as the PCL or the VET-R PTSD scale, with the CIDI PTSD module, it is the first to do so with a large, non-clinical sample.

Whether we consider a balance of sensitivity and specificity or LR+ and LR−, the optimal PCL threshold for a CIDI diagnosis of current PTSD is in the low to the mid-1930s. This is considerably lower than the cutoff score of 50 originally recommended (Weathers *et al.* 1993) based on data from treatment-seeking Vietnam combat Veterans. Subsequent research with a variety of populations indicates that lower cutoff scores in the 1930s may actually be typical in community samples (Boscarino, 2006; Foa

& International Society for Traumatic Stress Studies, 2009).

A recent review of 18 studies examined the accuracy of PCL scores in comparison to a PTSD diagnosis assessed with CAPS or SCID (McDonald & Calhoun, 2010). The optimal PCL score for diagnosis varied considerably from a low of 30 to a high of 60, but the study populations were heterogeneous and included both clinical and non-clinical samples. Those that studied non-treatment seeking Veterans reported PCL cutoff scores that were similar to what we found in the present study. Two studies of Veterans in primary care reported PCL cutoff scores of 30 (Lang *et al.* 2003) and 31 (Yeager *et al.* 2007). A study of 724 US

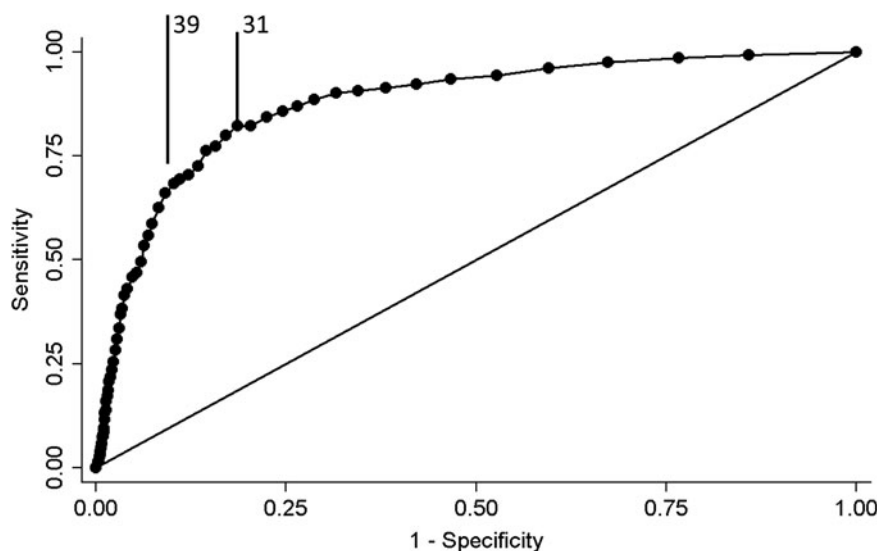


Fig. 2. ROC curve for VET-R PTSD scale.

Army soldiers post-deployment reported an optimal PCL cutoff score of 32 (Bliese et al. 2008). A small study of female primary care Veterans is the exception with a cutoff score of 50 (Lang & Stein, 2005).

The VET-R consists of Vietnam-Era Veteran males, so our results cannot be generalised to females or non-Veterans. Caution should also be used in making inferences about how the PCL would perform in younger Veterans from more recent conflicts and female Veterans; however, previous research has documented that the PCL performs well for these Veterans (Dobie et al. 2002; Bliese et al. 2008). We based our criterion diagnosis using a telephone administration of the CIDI PTSD module, which is not a clinically based assessment but an epidemiologic diagnostic tool. Thus, our results do not apply in clinical settings. Furthermore, it should be noted that the CIDI was not developed to be administered on the telephone, but rather as an in-person interview; therefore, the telephone version cannot be considered the 'gold standard' for such an instrument. Additionally, the time frames for the three instruments were different: 12 months for the CIDI, 1 month for the PCL and 6 months for the VET-R PTSD scale. Agreements may have been tighter with matching timeframes; nevertheless, we used the traditional timeframes for each instrument. Last, the order and context of the screening instruments may impact the psychometric performance of the instruments. In particular, there is the possibility that one instrument may have primed a subject to produce more thoughtful responses in the following instrument. In this study, the PCL always preceded the VET-R PTSD scale in the mail survey, and the CIDI was administered by telephone at a later date. Because the PTSD screeners were always presented in the same order, it is impossible to evaluate the impact of this potential bias. Also, the version of the PCL used in the study was meant for civilian use and does not reference military service, whereas the VET-R PTSD scale referenced military service directly. Even so, both performed slightly better among Veterans who served in Southeast Asia compared with those who did not serve in theatre. Strengths of this study include the use of a non-treatment seeking sample of older Vietnam-Era Veterans from all branches of service.

## Conclusions

The PCL and VET-R PTSD scales both demonstrate excellent psychometric performance compared with the CIDI PTSD module. Optimal PCL cutpoints for a PTSD diagnosis were lower than has been traditionally recommended but in line with current research with non-clinical populations.

## Acknowledgements

Dr Magruder had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The funding source was involved in the design and conduct of the study, and the interpretation, preparation, review and approval of the manuscript. The authors were responsible for the collection, management, analysis and interpretation of the data and for the preparation of the manuscript and its submission for publication. The authors gratefully acknowledge the continued cooperation and participation of the members of the VET-R: without their contribution this research would not have been possible. The authors would also like to thank the members of the Department of Veterans Affairs Cooperative Study #569 Group (in addition to the authors): I. Curtis, A. Ali, B. Majerczyk, B. Harp, K. Moore, A. Fox, M. Tsai, A. Mori, J. Sporleder, P. Terry, S. Kath, Seattle, WA; *Executive Committee*: S. Eisen, Washington, DC; A. Snodgrass, Albuquerque, NM. *Data Monitoring Committee*: J. Vasterling, Boston, MA; M. Stein, La Jolla, CA; B. Booth, Little Rock, AR; J. Westermeyer, Minneapolis, MN. *Planning Committee*: M. McFall, Seattle, WA; T. O'Leary, S. Eisen, Washington, DC; M. Smith, Palo Alto, CA; K. Swanson, Albuquerque, NM.

## Financial Support

The Cooperative Studies Program (CSP) of the US Department of Veterans Affairs' Office of Research and Development has provided financial support for Cooperative Study #569 and the development and maintenance of the Vietnam-Era Twin (VET) Registry. Dr Viola Vaccarino was supported in part by a National Institutes of Health award, K24 HL077506.

## Conflict of Interest

None.

## Ethical Standard

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

## Supplementary materials and methods

The Supplementary material referred to in this article can be found at <http://dx.doi.org/10.1017/S2045796014000365>.

## References

- American Psychiatric Association** (1987). *Diagnostic and Statistical Manual of Mental Disorders*, 3rd Revised edn. American Psychiatric Association: Washington, DC.
- American Psychiatric Association** (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC.
- Andrykowski M, Cordova M, Studts J, Miller T** (1998). Posttraumatic stress disorder after treatment for breast cancer: prevalence of diagnosis and use of the PTSD Checklist – Civilian Version (PCL-C) as a screening instrument. *Journal of Consulting and Clinical Psychology* **66**, 586–590.
- Blake D, Weathers F, Nagy L, Kaloupek D, Klauminzer G, Charney D, Keane T** (1990). A clinician rating scale for assessing current and lifetime PTSD: the CAPS-1. *Behavior Therapist* **13**, 187–188.
- Blake D, Weathers F, Nagy L, Kaloupek D, Gusman F, Charney D, Keane T** (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress* **8**, 75–90.
- Blanchard E, Jones-Alexander J, Buckley T, Forneris C** (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy* **34**, 669–673.
- Bliese P, Wright K, Adler A, Cabrera O, Castro C, Hoge C** (2008). Validating the primary care posttraumatic stress disorder screen and the posttraumatic stress disorder checklist with soldiers returning from combat. *Journal of Consulting and Clinical Psychology* **76**, 272–281.
- Boscarino J** (2006). Posttraumatic stress disorder and mortality among U.S. Army veterans 30 years after military service. *Annals of Epidemiology* **16**, 248–256.
- Boscarino J** (2008). A prospective study of PTSD and early-age heart disease mortality among Vietnam veterans: implications for surveillance and prevention. *Psychosomatic Medicine* **70**, 668–676.
- Dobie D, Kivlahan D, Maynard C, Bush K, McFall M, Epler A, Bradley K** (2002). Screening for post-traumatic stress disorder in female Veteran's Affairs patients: validation of the PTSD checklist. *General Hospital Psychiatry* **24**, 367–374.
- Eisen S, True W, Goldberg J, Henderson W, Robinette C** (1987). The Vietnam Era Twin (VET) Registry: method of construction. *Acta Geneticae Medicae et Gemellologiae (Roma)* **36**, 61–66.
- Elhai J, Gray M, Kashdan T, Franklin C** (2005). Which instruments are most commonly used to assess traumatic event exposure and posttraumatic effects?: a survey of traumatic stress professionals. *Journal of Traumatic Stress* **18**, 541–545.
- Foa E, International Society for Traumatic Stress Studies** (2009). *Effective Treatments for PTSD: Practice Guidelines from the International Society for Traumatic Stress Studies*, 2nd edn. Guilford Press: New York.
- Goldberg J, True W, Eisen S, Henderson W** (1990). A twin study of the effects of the Vietnam War on posttraumatic stress disorder. *Journal of the American Medical Association* **263**, 1227–1232.
- Goldberg J, Curran B, Vitek M, Henderson W, Boyko E** (2002). The Vietnam Era Twin Registry. *Twin Research* **5**, 476–481.
- Henderson W, Eisen S, Goldberg J, True W, Barnes J, Vitek M** (1990). The Vietnam Era Twin Registry: a resource for medical research. *Public Health Reports* **105**, 368–373.
- Hoge C, Castro C, Messer S, McGurk D, Cotting D, Koffman R** (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine* **351**, 13–22.
- Kessler R, Chiu W, Demler O, Merikangas K, Walters E** (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**, 617–627.
- Kraemer H** (1987). The methodological and statistical evaluation of medical tests: the dexamethasone suppression test in psychiatry. *Psychoneuroendocrinology* **12**, 411–427.
- Kraemer H, Kupfer D, Clarke D, Narrow W, Regier D** (2012). DSM-5: how reliable is reliable enough? *American Journal of Psychiatry* **169**, 13–15.
- Lang A, Stein M** (2005). An abbreviated PTSD checklist for use as a screening instrument in primary care. *Behaviour Research and Therapy* **43**, 585–594.
- Lang A, Laffaye C, Satz L, Dresselhaus T, Stein M** (2003). Sensitivity and specificity of the PTSD checklist in detecting PTSD in female veterans in primary care. *Journal of Traumatic Stress* **16**, 257–264.
- Manne S, Du Hamel K, Gallelli K, Sorgen K, Redd W** (1998). Posttraumatic stress disorder among mothers of pediatric cancer survivors: diagnosis, comorbidity, and utility of the PTSD checklist as a screening instrument. *Journal of Pediatric Psychology* **23**, 357–366.
- McDonald S, Calhoun P** (2010). The diagnostic accuracy of the PTSD checklist: a critical review. *Clinical Psychology Review* **30**, 976–987.
- Robins L, Helzer J, Croughan J, Ratcliff K** (1981). National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. *Archives of General Psychiatry* **38**, 381–389.
- Robins L, Wing J, Wittchen H, Helzer J, Babor T, Burke J, Farmer A, Jablenksi A, Pickens R, Regier D** (1988). The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry* **45**, 1069–1077.
- Roy-Byrne P, Arguelles L, Vitek M, Goldberg J, Keane T, True W, Pitman R** (2004). Persistence and change of PTSD symptomatology – a longitudinal co-twin control analysis of the Vietnam Era Twin Registry. *Social Psychiatry and Psychiatric Epidemiology* **39**, 681–685.
- Schlenger W, Caddell J, Ebert L, Jordan B, Rourke K, Wilson D, Thalji L, Dennis J, Fairbank J, Kulka R** (2002). Psychological reactions to terrorist attacks: findings from the National Study of Americans' Reactions to September 11. *Journal of the American Medical Association* **288**, 581–588.

- Schuster M, Stein B, Jaycox L, Collins R, Marshall G, Elliott M, Zhou A, Kanouse D, Morrison J, Berry S** (2001). A national survey of stress reactions after the September 11, 2001, terrorist attacks. *New England Journal of Medicine* **345**, 1507–1512.
- Spitzer R, Williams J, Gibbon M, First M** (1990). *Structured Clinical Interview for DSM-III-R, Patient Edition/Non-patient Edition, (SCID-P/SCID-NP)*. American Psychiatric Press: Washington, DC.
- StataCorp** (2009). *Stata Statistical Software*. StataCorp LP: College Station, TX.
- True W, Rice J, Eisen S, Heath A, Goldberg J, Lyons M, Nowak J** (1993). A twin study of genetic and environmental contributions to liability for posttraumatic stress symptoms. *Archives of General Psychiatry* **50**, 257–264.
- Weathers F, Litz B, Herman D, Huska J, Keane T** (1993). The PTSD checklist: reliability, validity, and diagnostic utility. In *Annual Meeting of the International Society for Traumatic Stress Studies*. San Antonio, TX.
- Weathers F, Ford J** (1996). Psychometric properties of the PTSD checklist (PCL-C, PCL-S, PCL-M, PCL-PR). In *BH Stamm, Editor, Measurement of stress, trauma, and adaptation*, Sidran Press, Lutherville, MD.
- Yeager D, Magruder K** (2013). PCL scoring rules for elderly Veteran Affairs outpatients. *American Journal of Geriatric Psychiatry* **13**, S1064–S7481. [Epub ahead of print].
- Yeager D, Magruder K, Knapp R, Nicholas J, Frueh B** (2007). Performance characteristics of the posttraumatic stress disorder checklist and SPAN in Veterans Affairs primary care settings. *General Hospital Psychiatry* **29**, 294–301.